MILESTONE 2

# Data Cleaning

* See the “audible-cleaned.xlsx” file for the cleaned dataset.
* **NOTE:** The “Stars” column has been converted to numeric data type, and instances of “Not rated yet” have been replaced with -1.0 (to indicate that no ratings exist for the record in question).

# Data Analysis

1. **Answer:** No (see the “Sheet2” sheet of the “finearts.xlsx” file for the relevant pivot table).

**Justification:** The numbers of women who are admitted or rejected from the College of Fine Arts are almost exactly the same as the numbers of men who are admitted or rejected from the College of Fine Arts.

1. **Answer:** Emergency Room (see the “Sheet2” sheet of the “hospital.xlsx” file for the relevant pivot table and table)

**Justification:** Although the overall percentage of patients who live (resp. die) is higher (resp. lower) for Chicago Hope than for Emergency Room, this difference is not significant as it is within two percentage points and the number of patients admitted is roughly the same (203 for Chicago Hope and 200 for Emergency Room). On the other hand, the percentages of patients who live (resp. die), within each of the high-risk and low-risk categories, are significantly higher (resp. lower) for Emergency Room than they are for Chicago Hope, even though Emergency Room has admitted a significantly higher number of high-risk patients, unlike Chicago Hope.

1. **Answer:** Yes (see the “Sheet2”, “Sheet3” and “Sheet4” sheets of the “dow.xlsx” file for the relevant pivot charts).

**Justification:** There are several years which have seen a disproportionately high average return (e.g., 1954 and 1976), as well as several years which have seen a disproportionately low average return (e.g. 1973 and 1974), when compared to rest of the yearly average returns. This is further complemented by the frequency and amplitude of consecutive monthly returns which are either positive or negative within the specific years in question, which form patterns present in only a small minority of years from 1947 to 1992.

1. **Answer:** See the “Sheet2” sheet of the “makeupdb.xlsx” file for the relevant pivot table. The function that always yield’s Jen’s lipstick sales is given (between quotation marks) by:

“*=GETPIVOTDATA("Sum of Dollars",$A$4,"Name","Jen","Product","lipstick")*”

1. **Answer:** 17

**Justification:** Let be the number of paid registrants which need to attend for me to break even. Then, given that there are speakers at the conference and that the conference fees entail per speaker and per conference participant for food and lodging costs, it follows that the number satisfies:

whereupon:

Therefore, 17 paid registrants need to attend for me to break even.